الإحصاءات غير المعلمية و الترتيبية

باستخدام MATLAB

أستاذ دكتور

محمد المهدي محمد على

أستاذ الرياضيات والإحصاء الاكتواري

رئيس قسم الإحصاء و الرياضيات والتأمين

ووكيل كلية التجارة – جامعة بورسعيد (سابقا)

2015/2014

مقدمة

من المعروف أنه أثناء دراستنا لعلم الإحصاء الوصفي أو مبادئ الإحصاء أن البيانات الإحصائية قد تكون في صورة كمية Quantitative أو كيفية (وصفية) Qualitative، ومن ناحية أخرى فإن البيانات الكيفية (أو الوصفية) قد تشمل نوعاً خاصاً من البيانات يكون في شكل ترتيبي، أو قد تسمى بالبيانات الترتيبية Ranked data.

وفي هذا الجزء من الكتاب نقدم للقارئ الأساليب الإحصائية المصممة أو المناسبة لإجراء اختبارات الفروض عندما تكون البيانات المتاحة في شكل ترتيبي، أو في حالة المجتمعات ذات البيانات الترتيبية.

ومن المعروف أنه عندما تكون البيانات غير كمية فإننا لا نستطيع حساب مقاييس المتوسطات أو التباينات، ومن ثم فإن الطرق أو الأساليب الإحصائية المقدمة في هذا الكتاب لا تمكننا من اختبار قيمة المتوسط أو التباين ولكنها تكون مناسبة فقط لاختبار صفات المجتمعات Characteristics of populations بدون الإشارة إلي معلمة أو معالم محددة. ولهذا السبب فإن هذه الأساليب أو الطرق الإحصائية تسمى بالأساليب أو الطرق غير المعلمية Nonparametric techniques ، والتي تختلف بدورها عن طرق الاستدلال الإحصائي السابق دراستها، والتي تسمى الطرق المعلمية Parametric techniques وعلى الرغم من أن الأساليب غير المعلمية التي سوف تناقش في هذا الجزء تستخدم أساساً لاختبار البيانات الترتيبية، فإنها تصلح للتطبيق في مجالات أخرى في حالة البيانات الكمية التي لا تتبع التوزيع الطبيعي.

فقد سبق أن ذكرنا أنه عند استخدام أساليب الاستدلال الإحصائي في حالة البيانات الكمية توجد عدة شروط يجب توافرها لكي نتمكن من استخدام توزيعات نظرية معينة. ففي حالة البيانات الكمية قد ذكرنا أنه عندما يكون التوزيع العيني للبيانات تابعاً للتوزيع (t) أو (F) فإن أحد الشروط اللازمة أو المطلوبة لاستخدام الطرق أو الاساليب الإحصائية في هذه الحالات أن المجتمع المسحوب منه العينة يجب أن يتبع التوزيع الطبيعي أو على الأقل يقترب منه.

فإذا كانت البيانات لا تتوزع طبيعياً أو على الأقل لا تقترب من التوزيع الطبيعي فإن نتائج تطبيق تلك الأساليب الأحصائية تكون غير واقعية أو مضللة، ومن ثم فإنه في مثل هذه الأحوال يفضل استخدام الأساليب غير المعلمية.

البيانات الترتيبية : Ranked Data

يمكن تصنيف العديد من البيانات الإحصائية في شكل بيانات ترتيبية. فمثلاً افترض أنك تجرى دراسة تسويقية عن نوع جديد من العصائر أو المشروبات، فقد تحتاج إلى سؤال بعض المستهلكين أو أخذ رأيهم عن مدى استساغة أو لذة طعم هذه المنتجات الجديدة، وفي هذه الحالة يتعين أخذ عينة من المستهلكين وتقديم هذا المنتج لهم لتذوقه ثم الإجابة على تساؤلاتك، وقد تكون الاختيارات المتاحة للإجابة على سؤالك هي أن يكون طعم المنتج ممتاز أو جيد أو مقبول أو غير مقبول (أي غير مستساغ).

ونظراً لأن الإجابات المتاحة تعتبر بمثابة بيانات كيفية غير كمية، فإننا نستطيع تحديد قيمة عددية Numerical value واحدة لكل إجابة من الإجابات الأربعة الممكنة حتى يمكن تبويب وتصنيف هذه البيانات، ولكن إذا تم تحديد قيم عددية بطريقة حكمية

أو عشوائية تماماً، فقد لا نستطيع معرفة التسلسل الترتيبي للإجابات، أي أننا قد نفقد بعض المعلومات اللازمة لترتيب الإجابات في شكل متسلسل أو هرمي Hierarchical order، وبالتالي يجب أن نحدد قيماً عددية متتالية للإجابات المتتالية حسب أهميتها للدراسة لكي نحافظ علي النظام العددي اللازم لترتيب الإجابات. فمثلاً قد نحدد قيماً عددية للإجابات المتتالية علي النحو التالي:

طعم ممتاز= 4 ، طعم جيد= 3

طعم مقبول= 2 ، طعم غير مقبول= 1

أو قد نضع قيماً عددية أخرى مثل: (10 ,8 ,7 ,1)، أو أي قيم أخرى بحيث نحافظ على إعطاء قيمة أكبر للإجابة الأفضل وهكذا.

وبوضع هذه القيم العددية على النحو السابق ذكره، نكون قد حولنا البيانات من وصفية إلي كمية من حيث الشكل وليس المضمون، حيث لا يمكن القول بأن هذه البيانات أصبحت كمية، وبالتالي تعالج علي أساس الأساليب الإحصائية المستخدمة في حالة البيانات الكمية.

فإذا افترضنا أن عينة عشوائية مكونة من 12 من المستهلكين تم سؤالهم عن مدى جودة المنتج المذكور، وكانت الإجابات تأخذ الترتيب المشار إليه سابقاً: ممتاز، جيد، مقبول، ردئ. وإذا كانت نتيجة العينة مثلاً هي: 8 مشاهدات ممتاز و4 ردئ، فإن متوسط العينة باستخدام النظام العددي الأول يكون كما يلي:

$$\bar{x} = \frac{1\times4+4\times8}{12} = 3$$

أي أن قيمة المتوسط تناظر جيد، بينما في النظام العددي الثاني نجد أن المتوسط يكون كما يلي:

$$\bar{x} = \frac{1\times 4+10\times 8}{12} = 7$$

أي أن قيمة المتوسط تناظر مقبول، وعلي ذلك نجد أن المتوسط المحسوب باستخدام النظام العددي الأول يختلف عن ذلك المحسوب باستخدام النظام الثاني. فإذا تعاملنا مع هذه البيانات علي أنها بيانات كمية، فإن تفسير نتائج الإجابات سوف يعتمد علي قيم النظام العددي الحكمية التي تم وضعها للإجابات المختلفة.

ولكي نتحاشى مشكلة معالجة البيانات علي أنها كمية وفقد بعض المعلومات عن تسلسل الإجابات، أو معالجتها علي أنها كمية ثم حساب معلمات مضللة أو لا معنى لها مثل المتوسط أو التباين، فقد استخدم الإحصائيون أساليب إحصائية غير معلمية في مثل هذه الحالات.

وسوف نرى فيما بعد أن الاختبارات غير المعلمية تستخدم أسلوباً أو إجراءً ترتيبياً كجزء رقمي أو عددي في الحسابات.

فأثناء دراستنا للإحصاء الوصفي، وخاصة مقاييس النزعة المركزية (المتوسطات) تعلمنا كيف نستخدم مقياس الوسيط كأحد مقاييس النزعة المركزية أو الموضع.

وكما نعلم فإن الوسيط Median يحسب عن طريق ترتيب البيانات (المشاهدات) تصاعديا أو تنازليا، ثم إيجاد القيمة التي تقع في المنتصف، أي التي يقل عنها في الترتيب نصف البيانات ويزيد عليها في الترتيب النصف الاخر، فيقال أن هذه

القيمة هي القيمة الوسيطية أو وسيط البيانات. وعلي ذلك فإن مقياس النزعة المركزية للبيانات الترتيبية هو الوسيط، أي هو الإحصاء الناتج عن عملية الترتيب.

وعلى ذلك نستطيع تصنيف البيانات الإحصائية عموماً إلى ثلاثة أنواع من البيانات هى: البيانات الكمية والبيانات الترتيبية والبيانات الكيفية أو النوعية. وهذه الأنواع الثلاثة من البيانات يمكن ترتيبها وفق نظام معين للأغراض الحسابية التي تسمح بذلك.

فالبيانات الكيفية أو النوعية تعد الأقل ترتيباً حيث لا تسمح بإجراء أية عمليات حسابية فيما عدا النسب المئوية، ثم تأتي البيانات الترتيبيه في المرتبة الثانية نظراً لأننا نستطيع حساب بعض المقاييس لهذه البيانات علي أساس الترتيب مثل الوسيط والربيعات وغيرها من المقاييس المكانية، أما البيانات الكمية فتصنف في المرتبة أو الدرجة الأولى نظراً لأننا نستطيع التعامل معها حسابياً في كل الأغراض الإحصائية.

و في هذا الكتاب سوف نستعرض اختبارات الفروض فى حالة الإحصاءات اللامعلمية أو غير المعلمية كما يفضل تسميتها بالإضافة إلى إختبارات x^2 و معامل إرتباط الرتب . ومن أهم هذه الاختبارات اختبار الإشارة، واختبار الدورات، واختبار ويلكوكسون للعينات المستقلة والعينات المرتبطة، واختبار كروسكال– واليز للتجارب تامة العشوائية، واختبار فريدمان للمجموعات أو البلوكات كما يلى :

1– اختبار الإشارة لعينة واحدة.

2– اختبار الإشارة لعينتين مرتبطتين.

3- معامل ارتباط الرتب واختبار الدورات.

4- اختبار ويلكوكسن لمجموع الرتب لعينتين مستقلتين.

5- اختبار ويلكوكسن لمجموع الرتب والإشارة لعينتين مرتبطتين.

6- اختبار كروسكال-واليز لعينتين مستقلتين أو أكثر.

7- اختبار فريدمان لعينتين غير مستقلتين أو أكثر.

8- إختبار الدورات .

9- إختبارات x^2 .

10- معامل إرتباط الرتب .

11- إختبارات الإعتدالية (كولمجروف- سميرنوف ، ليليفورس).

قائمة المراجع

1- David F.G and patrick W. S Business Statistics , A decision - Making Approach , A bell & Howell Co. , Columbus , 1981.

2- David R. Anderson et al , Statistics for Business and Economics , 11 th edition , South Western , Cengage Learning , 2011 .

3- Douglas , C.M and George C-R, Applied Statistics and Probability for Engineers , Second Edition , John. Wiley & Sons , Inc.,N.Y., 1999.

4- Donald Waters , Quantitative Methods For Business , Fourth Edition , F.T, Prentice Hall , 2001 .

5- John W.B., Donald R.S., Methods of Finite Mathematics, John. Wiley & Sons, N.Y., 1989.

6- Keller, Warrack and Bartel , Statistics For Management and Economics , Duxbury , Press , Belmont , 1994.

7- Lawrence B. Morse, Statistics for Business and Economics, Harper Collins College Publishers,1996.

8- Mendenhall , Scheaffer and Wackerly , Mathematical Statistics With Application , Duxbury Press , Boston 1986.

9- Mizrahi , Sullivan Mathematical Statistics , An Applied Approach , John Wiley & Sons , Inc N.Y.,1996.

10- Park J.E, James S.F and Chi-y. lin , Applied Managerial Statistics , Prentice – Hall , Inc. , N.J., 1982.

11- Prem S. Mann , Statistics for Business & Economics , John Wiley & Sons , inc. , N.Y., 1995.

12- Prem S.Mann, Introductory Statistics , Sixth Edition , John wiley & Sons , Inc.,2007 .

13- Walpole Myers and Myers ,Probability and Statistics for Engineers and Scientists , Hall Inter-national ,Inc, N. Jersey, 1998.

الفهرس

للحصول على نسخة الكتاب كاملا

المكتبة الأكاديمية 121 شارع التحرير الدقى ، القاهرة ، مصر.